

DOCUMENT RESUME

ED 069 715

TM 002 167

AUTHOR Karma, Kai  
TITLE Investigations into the Instructional Process: V. Experiences with the Bellack Classification System.  
INSTITUTION Helsinki Univ. (Finland). Inst. of Education.  
PUB DATE Jan 72  
NOTE 19p.; Research Bulletin No. 30  
EDRS PRICE MF-\$0.65 HC-\$3.29  
DESCRIPTORS \*Classification; Classroom Communication; \*Codification; Grade 3; Grade 4; Grouping (Instructional Purposes); Instruction; Instructional Programs; \*Student Teacher Relationship; Teacher Role; Teaching Methods; Testing; \*Test Interpretation; \*Test Reliability

ABSTRACT

Video- and sound-tape recordings of 96 lessons in one Finnish school were analyzed using the Bellack Classification System. The lessons were composed of Finnish, arithmetic, civics, drawing, religion, and music for Grades 3 and 4. During the coding process, it was found that the Bellack system was not very suitable for describing instruction at the elementary level. Consequently, the system was modified. Results are presented for the following categories: pedagogical moves; logical, instructional, rating, and extralogical substantive meanings; and cycles. The results indicated that the teacher's role that emerged was rather similar to Bellack's description. However, differences were found in the results for the pupil's role. (JS)

ED 069715



## RESEARCH BULLETIN

Institute of Education  
University of Helsinki

Head:  
Matti Koskenniemi  
Professor of Education

SF - 00100 Helsinki 10  
Finland

TM 002 167

No. 30

January 1972

Kai Karma

INVESTIGATIONS INTO THE INSTRUCTIONAL PROCESS  
V. Experiences with the Bellack Classification System



Kai Karma

INVESTIGATIONS INTO THE INSTRUCTIONAL PROCESS

V. Experiences with the Bellack Classification  
System

Institute of Education  
University of Helsinki  
1972

Contents:

|                                                     |    |
|-----------------------------------------------------|----|
| 1. Introduction .....                               | 2  |
| 2. Reliability .....                                | 4  |
| 3. Results .....                                    | 8  |
| 3.1. Pedagogical Moves .....                        | 8  |
| 3.2. Logical Categories .....                       | 9  |
| 3.3. Instructional Categories .....                 | 10 |
| 3.4. Rating and Extralogical Categories .....       | 11 |
| 3.5. Substantive Meanings .....                     | 11 |
| 3.6. Cycles .....                                   | 11 |
| 4. Pedagogical Roles in the Light of the Results .. | 13 |
| 4.1. Teacher's Role .....                           | 13 |
| 4.2. Pupil's Role .....                             | 14 |
| References .....                                    | 16 |

## Investigations into the Instructional Process

### V Experiences with the Bellack Classification System

#### 1. Introduction

At the University of Helsinki Institute of Education, a research group is engaged in the analysis of the instructional process under the direction of Professor Matti Koskenniemi<sup>1)</sup>. Since 1967 video- and sound-tape recordings of lessons at one elementary school class grades 3 and 4 have been made for the purpose. The tapes have been analyzed employing the Bales, Bellack and Flanders classification systems. The material dealt with in this paper was recorded between fall 1967 and spring 1969. As the composition of the class remained virtually unchanged throughout this period of time and as there were only two teachers, this investigation has to be regarded as a kind of case study. The material consisted of 96 lessons, analyzed by the Bellack system, the school subjects covered being Finnish, arithmetic, civics, drawing, religion, and music. The codings were carried out directly from the tapes, and thus the interaction was not written down.

Even when coding was being practiced, it was found that the Bellack system was not very well suitable as such for the description of instruction at the elementary level. In

---

<sup>1)</sup> The other members of the group are:  
Anna-Kaarina Falck, Pentti Holopainen, Kai Karma,  
Erkki Komulainen and Risto Skoglund.

several cases the categories were too many in number, it was difficult for the coders to perceive so slight differences unanimously, and the frequencies in many categories remained very low. Therefore the system was simplified. A brief account is given below of the system in the form it was applied here.

Pedagogical moves are identical with those in the original system.

Logical meanings are coded into a single category. The substantive-logical and instructional-logical categories proved difficult to distinguish from each other, and in most cases their codings were the same nonetheless. DEF, DED, DEC and INT were combined because of low frequencies and slight differences in meaning. The combined category was designated DEF. For similar reasons, categories OPN and JUS were combined into a single category, termed OPN. Thus the main category of logical meanings comprises four subcategories: DEF, FAC, XPL and OPN.

Instructional meanings. LOG and ACT were omitted. LOG can be coded into other categories (DEF, LAM, STA). There was a tendency for ACT to become a category to which all utterances difficult of coding were assigned. A category DIS was separated from PRC to signify the maintenance of discipline. Thus the main category of instructional meanings included the following subcategories: ASG, MAT, PRC, PER, STA, ACV, ACC, ACP, ACE, DIS, LAM.

Rating and extralogical categories. Evaluative and extralogical meanings were separated from the instructional main category to form a group on their own. NEG and NAD were combined into one, because they were difficult to distinguish. AON and PON were combined with DIR, mainly because of low frequencies. Thus the rating and extralogical categories were POS, ADM, RPT, QAL, NEG, DIR, PRF.

Substantive meanings. These were not classified. Only their occurrence in the utterance was coded. An utterance was considered to involve substantive meanings when it dealt with matters associated with the theme of the ongoing lesson.

## 2. Reliability

Disagreement between two coders can, in principle, arise in two different ways:

a) Interaction is divided into units differently. One coder may regard as a separate move a part of an utterance which the other combines with the preceding move or, which often happens when coding is done directly from the tape, the coders hear differently and omit parts of interaction that they consider to be meaningless.

b) The division into units is identical, but different codings are given to the units.

In determining reliability, those parts of interaction that are registered only by one coder and not by both have been omitted. These have accounted for some 10 - 15 % of the total moves. If the intention is to determine the systematic differences between the coders, it suffices to compare the total markings in each category with each other. The reliability index will then be the percentage ratio of the smaller to the larger number. Ordinarily, reliability has been computed by comparing through cross-tabulation the codings by the two coders move by move. The reliability index for any one category will thereby be the percentage proportion of identical codings of all the codings in the category.

The reliability coefficients expressed as percentages have the drawback that they provide a misleading picture of agreement between the coders in cases where there is little agreement. Even when there is no agreement, i.e., when the codings are distributed among the categories at random, per-

centage indexes are definitely positive. Also, for one accustomed to the use of correlative indexes it is difficult to obtain a correct picture of percentage indexes.

For these reasons it is suggested that reliability be computed as phi correlation coefficients. When the codings have been cross-tabulated, it is possible to find out both the number of cases where the coders' codings are similar and the number of cases where they are different. These can be written in a  $2 \times 2$  table by dividing both the similar and different codings evenly on their own diagonals. The halves of the diagonals will then be the same ( $a = d$  and  $b = c$ ) and the marginal sums will be equal ( $a+b = a+c = b+d = c+d$ ). Hence, the formula of the coefficient can be simplified:

$$\frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} = \frac{a^2 - b^2}{\sqrt{(a+b)^4}} = \frac{a^2 - b^2}{(a+b)^2}$$

In the case of move types and logical categories we have a  $4 \times 4$  table, where the number of off-diagonal cases is three times that of diagonal cases if we have a chance distribution, i.e., zero correlation. In order that each cell of the  $2 \times 2$  table should include the same number of cases (= zero correlation), the number of off-diagonal cases must be divided by three. When written in terms of the frequencies, the formula then becomes:

$$\frac{f_o^2 - \left(\frac{f_u}{3}\right)^2}{\left(f_o + \frac{f_u}{3}\right)^2} \quad \begin{array}{l} f_o = \text{number of diagonal cases} \\ f_u = \text{number of off-diagonal cases.} \end{array}$$

When use is made of Bellack's original classification, we have a  $4 \times 4$  table only for the move types, whereas the tables for the other main categories will be larger. For the sake of

uniformity, it is suggested that the coefficient 3 be used with all of the main categories. The reliability coefficient will then be the same if the percentage proportion of similar codings is the same. Moreover, a case produced truly by chance is difficult to define, except for move types, since the rest of the categories do not exclude one another equally clearly.

In order to make the reliabilities for the major categories and subcategories mutually comparable,  $f_u$  has to be divided by two in computing the index for a subcategory or, which is often more expedient,  $f_o$  has to be multiplied by two. This is necessary in order not to double-count the different codings.

Example:

|    |    |    |    |
|----|----|----|----|
| -  | 5  | -  | 30 |
| 5  | -  | 10 | -  |
| 5  | 10 | -  | 5  |
| 20 | -  | 10 | -  |

The reliability of the entire main category:

$$\frac{70^2 - \left(\frac{30}{3}\right)^2}{70 - \left(\frac{30}{3}\right)^2} = \frac{4900 - 100}{6400} = .75$$

Reliability of subcategory A:

$$\frac{(2 \times 20)^2 - \left(\frac{20}{3}\right)^2}{2 \times 20 + \left(\frac{20}{3}\right)^2} = \frac{1600 - 44}{2171} = .71$$

The consistency of any one coder has usually been rather good. In one case, where the interval between the codings was

approximately half a year, the following reliabilities (phi coefficients) were obtained: move types, .95; logical categories, .92; instructional categories, .72; extralogical and rating categories, .78; substantive meanings, .86.

The differences between two coders often tend to become systematic. To avoid this, obscure points should be repeatedly controlled, example codings should be collected, etc. Typical between-coder reliability coefficients: move types, .80 - .90; logical categories, .60 - .75; instructional categories, .65 - .85; extralogical and rating categories, .70 - .80; substantive meanings, .70 - .80.

SOL and RES are the most clear-cut categories and are not in need of any supplementary definition. By contrast, REA and, in particular, STR are obscure in many instances, mainly in the sense that they are confounded with each other. At the grade-level dealt with in this study, pupils are very spontaneous, and this renders coding difficult. In cases where REA is associated with STA, i.e., where it is a comment on the preceding utterance, the interpretation is clear. Also, agreement is good regarding STR moves that are clearly forwards directed and initiate cycles in which the next move is either a SOL or REA move. However, among pupils' utterances, in particular, there are many that are not closely connected with the ongoing interaction; they are neither forwards nor backwards directed, nor do they usually result in changes in the course of interaction. They include expressions of ideas that suddenly occur to the pupils, answers meant to be amusing and usually addressed to other pupils, pupils' talk to one another not related to the topic discussed, etc. The reliability of STR might improve if not only utterances directed forwards but also those that are not clearly backwards-directed were included in it.

Of the logical categories, FAC and XPL are easily confounded. The difference between the two is often obscure and there is room for various interpretations; thus, a more de-

tailed definition seems to be called for.

Of the instructional categories, PRC, MAT and ASG are easily confounded.

Particularly two evaluative categories, POS and ADM, are easily confounded. The information provided by them is interesting, however, and thus it would be desirable to have their reliability improved without combining them in one. Coding is sometimes made difficult by the fact that for such neutral reactions as "mm" or "uhuh" - which can be interpreted as meaning "I understand" or "I heard what you said" - there are not symbols of their own. Such reactions could perhaps be assigned to RPT, which is also otherwise rather neutral in meaning. Another possibility would be to agree that they will be coded REA/FAC/ACC.

### 3. Results

#### 3.1. Pedagogical Moves

STR move. Here the results differ most clearly from Bellack's results. On an average, 17.5 % of all the moves in our material were STR moves, and almost a half of these moves (43 %) are accounted for by pupils. In Bellack's material, no more than 12 % of STR moves are pupils' moves. Even when the fact, indicated in discussing reliability, is taken into consideration that pupils' STR moves include talk not closely related to the course of the interaction, the tendency is clear. The frequency of occurrence of STR was highest in the civics lessons: the proportion of pupils' moves falling within this category being as large as 40 %. In those lessons STR is often connected with PER.

SOL move. A similar tendency, though less pronounced, is perceptible in SOL moves. They account for 36 % of the total interaction, and 24 % of them are pupils' moves (Bellack: 14 %).

Of the total moves by pupils, 22 % are SOL moves (Bellack: 11.3 %), the corresponding figure for teachers being 45 %.

RES move. In contrast to SOL and STR moves, which were more frequent in pupils' interaction here than in Bellack's results, RES moves were less frequent. Only 38 % of the total moves of pupils were RES moves (Bellack: 65.4 %). Of the total RES moves, pupils accounted for 80 percent (Bellack: 88 %).

REA move. Of the total moves, 27.4 % were REA moves. Here, too, the proportion accounted for by pupils was significantly larger than in Bellack's material, or 31 % (Bellack: 19 %). The share of REA-P of the total moves was 1.3 % with pupils and 2.6 % with teachers. REA underlined moves are mainly teachers' moves; here their share of all teacher's moves was 1.3 %. Less than 0.5 % of the total moves were of the REA<sup>+</sup> type, and a large majority of these moves were teachers' moves.

### 3.2. Logical Categories

The average proportions of the total moves accounted for by the various logical categories were as follows: DEF, 2.5 %; FAC, 60.9 %; XPL, 8.4 %; and OPN, 2.6 %. No logical meanings were involved in 25 % of the moves. The moves involving logical meanings were distributed approximately evenly between pupils and teachers, except XPL moves, which were more frequent with teachers (68.4 %). The present results differed very clearly from Bellack's. Compared with Bellack's results, FAC was almost three times as high, at the expense of the other logical categories. It would seem that pedagogically meaningful is just the ratio that FAC bears to the other logical categories. The explanation presenting itself first is that the difference between these and Bellack's results is

ascribable to the grade-level. It would seem natural for instruction to include more "hard facts" on the elementary level than on the more advanced levels. It should be kept in mind, however, that the reliability of the XPL category, in particular, would be in need of improvement.

### 3.3. Instructional Categories

An average of 66 % of the total moves has involved instructional meanings. The average proportions of the total moves accounted for by the various categories have been as follows: STA 14.9, ASG 15.9, MAT 6.3, PRC 5.7, DIS 1.6, PER 3.8, ACV 8.4, ACC 4.5, ACP 3.6, ACE 0.8, LAM 0.6. Both the total frequency (which exceeded the 46.9 % reported by Bellack) and the percentages for the various categories indicate that utterances related to the learning and instruction process were more frequent here than in Bellack's material. Correspondingly, utterances related to the school subject concerned were relatively less frequent here. The percentages for ASG, MAT, PRC + DIS and ACP were definitely higher than with Bellack. On the other hand, STA and ACV, which generally have to do with the school subject itself, were less frequent here. The most natural explanation again seems to lie in the grade-level. It should also be taken into account that, in the present study, the material included subjects (such as drawing) in which the performance of tasks and the discussion concerning them give rise to such a tendency, irrespective of the grade-level. The frequent use of PER is obviously typical of the elementary level in particular. STA, PRC, DIS, ACV, ACP and LAM were definitely more frequent with teachers than pupils, whereas the opposite was the case with PER and ACE.

### 3.4. Rating and Extralogical Categories

An average of 35 % of the total moves fell within rating and extralogical categories. Such categories were involved in almost a half of teachers' moves (46.3 %), the corresponding figure for pupils being 17.3 %. The average proportions of total moves accounted for by the various categories were as follows: DIR 2.6, PRF 15.3, POS 1.0, ADM 7.7, RPT 2.8, QAL 1.8 and NEG 2.4 %. Teachers used the following definitely more than pupils: PRF (93.5 % of total), POS (95.8 %), ADM (92.3 %), QAL (95.0 %). By contrast, compared with teachers, pupils used DIR (98.5 %) and NEG (53.9 %) more frequently. The percentages for DIR and PRF, which were higher than those reported by Bellack, seem to suggest that the need for advice and guidance is greater at the elementary level, but an effect has also certainly been contributed by the large number of work tasks, referred to above, and by the discussion of these tasks. The pupils' spontaneity is evidenced particularly clearly by NEG. Pupils are very eager to point out mistakes, committed by other pupils and also by teachers. On the other hand, teachers tend to avoid using NEG; instead, they try to show the incorrectness of, e.g., an answer by milder means.

### 3.5. Substantive Meanings

Less than half of the total moves (43.6 %) involve substantive meanings. These are distributed approximately evenly between teachers and pupils.

### 3.6. Cycles

Cycles have not yet been counted from the total material, the following information on them being based on studies made

of a part of it only. The cycles used here are of a formal type. The liveliness of lessons and spontaneity of pupils make it often difficult to form cycles. For example, the STR-coded exclamations give automatically rise to an excessive number of STR-initiated cycles in the records if the cycles are formed afterwards, on the basis of the codings alone and without paying attention to the content of the utterances and to the way they are actually intended to be related with one another. Also, cycles that "do not exist" often tend to emerge. In order for the cycles to follow the actual logical course of the discussion, i.e., in order to improve their validity, we have begun to indicate them in the records even in connection with the coding. Such cycles are also not temporal: frequently they do not consist of temporally consecutive utterances. The decisive factor is the speaker's intention, in so far as this can be perceived during coding. Inadequate utterances will thereby not interrupt the cycle, nor will they be associated with cycles to which they are logically unrelated, but will, instead, form cycles on their own and often consist of a single utterance.

Cycles of the most frequent occurrence were 13 (SOL), 14 (SOL-RES), 1 (STR), 2 (STR-SOL), 3 (STR-REA) and 18 (SOL-RES-REA). There was a lot of variation in the distribution of the cycles, depending on school subject and teacher. The following average percentages were found: cycle 13: 20-28 %, cycle 14: 10-14 %, cycle 1: 6-17 %, cycle 2: 6-9 %, cycle 3: 8-17 % and cycle 18: 10-17 %. The frequent occurrence of STR moves increased the frequency of STR-initiated cycles. The high percentage of cycle 13 (SOL) may have been due to the grade-level; this cycle was often connected with categories PRF, DIS, MAT, PRC and ACP, i.e., with classroom activities not related to the school subject concerned. The liveliness of interaction is clearly reflected by the activity index, which was generally 6 - 7 cycles/min. The cycles contain an average of about 2 moves.

#### 4. Pedagogical Roles in the Light of the Results

##### 4.1. Teacher's Role

The picture of a teacher's role that emerged here was rather similar to Bellack's description. The teacher is clearly the most active person in the class, about 60 % of the total moves being made by her. The teacher's main function is soliciting: she tries to have the pupils' respond either verbally or non-verbally. The response is quite frequently a non-verbal one, as the number of pupils' RES moves only equals two-thirds of teachers' SOL moves. In some cases it is also necessary for the teacher to make more than one SOL move in order to obtain a single response. It is the teacher who principally employs SOL moves, about three-fourths of these being made by her.

The teacher's function next in importance is reacting to pupils' responses. One third of teachers' moves are REA moves. What the teacher reacts to is usually verbal action but it may sometimes also be non-verbal action. As a rule, the teacher reacts to an immediately preceding utterance, whereas summarizing reactions are comparatively infrequent. The teacher's evaluative reactions are usually mildly approving or neutral, whereas she endeavours to avoid clearly negative reactions. Reacting is mainly up to the teacher, about two-thirds of the REA moves being made by her.

The longest STR moves and the STR moves with the greatest effect on the course of the game are characteristics of the teacher's role. In most cases the teacher makes use of a STR move with the object of elucidating the matter about which she is going to ask a question or of indicating that she will proceed to consider a new topic. Thus, the teacher's STR move is usually followed by a SOL move.

The teacher responds comparatively rarely. What is concerned in these rare instances is usually a SOL move on the part of a pupil asking for advice.

#### 4.2. Pupil's Role

As already indicated in several contexts, the pupils were here clearly more spontaneous than Bellack's pupils. This does not mean, however, that the number of pupil's moves would have been greater. The distribution of moves between teachers and pupils was approximately the same as with Bellack: about 60 % of the total moves were made by the teacher.

As with Bellack, the pupil's main function was to respond. Nevertheless, the RES move occupied not nearly as prominent a position. Only 38 % of pupils' moves fell within this category. Pupils generally respond in one way or another to the stimuli provided by the teacher, but a majority of their responses are non-verbal, and thus they are not reflected in the codings.

In clear contrast to Bellack's case, the pupils in our study used the STR move notably often. Despite the fact that pupils' STR moves were usually shorter than those of teachers and that the cycles started with such moves contained a lower average number of moves than the cycles initiated by teachers' STR moves, pupils spontaneously influenced the course of interaction to a very great extent. Even if inadequate comments and remarks were disregarded, a great many utterances would remain that have been taken into consideration and that have a bearing on how the interaction continues. Usual are, for instance, accounts given by pupils of their own experiences that the matters dealt with during the lesson have made them recollect. Also, the pupils not infrequently express their own views and opinions even when the teacher does not ask them. Thus, in this material, the "don't structure" rule does not apply. A pupil's STR move is often followed by a REA move.

About a fifth of pupils' moves are SOL moves. A majority of the SOL moves addressed to the teacher appertain to instructional matters and are often related to DIR. Questions

related to substantive matters are asked by the pupils particularly when they have to perform lengthy tasks forming an integral whole and feel uncertain about how to continue ("Can this be colored red?" etc.). SOL moves addressed to other pupils are comparatively usual, and they may be related both to instructional and substantive matters.

Although a majority of the evaluative reactions are accounted for by the teacher, pupils also resort to such reactions notably often. Pupils react particularly spontaneously and eagerly to mistakes committed by others. Unlike teachers, the pupils do not generally attempt to soften down their negative reactions but express them quickly and sharply. Thus, the "don't react evaluatively" rule does not apply in the present material.

References:

BELLACK, A. A. & DAVITZ, J. R. 1963. The Language of the Classroom. Meanings Communicated in High School Teaching. New York: Institute of Psychological Research, Teachers College, Columbia University.

BELLACK, A. A. & KLIEBARD, H. & AL. 1966. The Language of the Classroom. New York: Teachers College Press.

**Previous Issues in this Series, Available from the Institute**

- No. 17 The Finnish Senior Secondary Research Project I. General Presentation of the Project and Its Original Subjects by ANNA-LIISA SYSIHARJU. Sept., 1967. 31 pp.
- No. 18 The Finnish Senior Secondary Research Project II. Departure from Junior Secondary and Transfer to Senior Secondary by ANNA-LIISA SYSIHARJU. Oct., 1967. 52 pp. Out of print.
- No. 19 Relationships between Absences from School and Certain Factors Related to School. A Study on Elementary School Grades Three to Six by PERTTI KANSANEN. Nov., 1967. 10 pp.
- No. 20 On the Anxiety Associated with School Attendance and the Grammar School Entrance Examination by EEVA PATJAS. May, 1968. 16 pp.
- No. 21 School Achievement and Personality. I: Design and Hypotheses, IV: Results and Discussion by ERKKI A. NISKANEN. Oct., 1968. 51 pp.
- No. 22 School Achievement and Personality. II: Operations at the Variable Level by ERKKI A. NISKANEN. Oct., 1968. 124 pp.
- No. 23 School Achievement and Personality. III: Operations at the Factor Level by ERKKI A. NISKANEN. Oct., 1968. 39 pp.
- No. 24 The Effectiveness of Two Methods of Teaching English as a Foreign Language in Some Finnish Secondary Schools by DANIEL J. CASEY. Nov., 1968. 37 pp.
- No. 25 The Learning of Elementary Algebra. An Empirical Investigation of the Results of Learning in a Simplified School Learning System by PAAVO MALINEN. Jan., 1969. 85 pp.
- No. 26 Investigations into the Instructional Process. I. Some Methodological Problems by MATTI KOSKENNIEMI and ERKKI KOMULAINEN in collaboration with Anna-Kaarina Falck and Pentti Holopainen. Oct., 1969. 35 pp.
- No. 27 Investigations into the Instructional Process. II. Objectivity of Coding in a Modified Flanders Interaction Analysis by ERKKI KOMULAINEN. Dec., 1970. 26 pp.
- No. 28 Investigations into the Instructional Process. III. P-technique Treatment of Observational Data by ERKKI KOMULAINEN. Jan., 1971. 41 pp.
- No. 29 Investigations into the Instructional Process. IV. Teaching as a Stochastic Process by ERKKI KOMULAINEN. Feb., 1971. 24 pp.

Distribution by EDUCA r.y., Fabianinkatu 28, SF - 00100 Helsinki 10